A Novel Deep Learning Framework For Rainfall Prediction In Weather Forecasting

[1]Dr.S. Soundararajan, [2]Dr. Chinnuswamy , [3]L. Maria Michael Visuwasam, [4]Dr. V.P Gladis Pushparathi, [5]Dr M.Sudha

[1]Professor
Department of Computer Science and Engineering
Velammal Institute of Technology, Panchetti, Chennai

[2]Professor
Department of Electronics and Communication Engineering
Velammal Institute of Technology, Panchetti , Chennai

[3]Assistant Professor-Selection Grade
Department of Computer Science and Engineering
Rajalakshmi Institute of Technology, Chennai

[4]Associate Professor
Department of Computer Science and Engineering
Velammal Institute of Technology, Panchetti, Chennai

[5]Assistant Professor –III
SRC, SASTRA Deemed to be University
Kumbakonam

ABSTRACT: Precipitation data may be used to evaluate important water supplies, drainage, ecosystems, and hydrology. For these purposes, data-driven model predictions using deep learning algorithms are promising. Rainfall is one of the most important sources of freshwater for all living things on the planet. The rainfall prediction model demonstrates how various climatological variables influence rainfall amounts. By allowing self-learning data labels, Deep Learning has recently allowed the development of a data-driven model for a time series dataset. As data availability and computational power have increased, deep neural network architectures have made significant progress in predicting problems across multiple domains. As data availability and computational power have increased in recent years, deep neural network frameworks have made significant progress in predicting problems across multiple domains. In this paper, a rainfall prediction system based on deep learning convolutional neural networks is proposed to provide insights into changes in environment and atmosphere variables such as precipitation, temperature, and humidity.

1. INTRODUCTION

Rainfall is now widely regarded as one of the primary causes of the majority of major incidents around the world. In Australia, agriculture is entirely reliant on rainfall, and it is one of the most important factors in deciding the country's economy. Apart from that, it is important to understand the amount of rainfall in coastal areas around the world. Rainfall forecasting should be done ahead of time to instal a rainwater harvester in some of the areas where there is a water shortage. Rainfall was forecasted primarily using satellite imagery, ground monitoring stations, and hot air balloons.

Rainfall forecasting is a source of concern for state governments, corporations, risk management organisations, and the scientific community. Rainfall is a climatic factor that affects many aspects of life, including agriculture, development, electricity generation, forestry, and hospitality, to name a few. Rainfall forecasting is crucial in this regard because it is the variable most closely linked to natural disasters such as landslides, floods, mass movements, and avalanches. These activities have had a long-term cultural effect. As a result, knowing how to forecast rainfall helps you to take preventative and mitigation measures against natural disasters Aguasca et al.[2019].

Radar imaging equipment is also available, but it is seldom used due to the large amount of image data it produces. The latest rainfall forecasting from the Australian Central Weather Bureau is added to global requirements, which include data from weather balloons, land meteorological stations, and satellite space exploration collected twice daily. A 12-hour temporal interval is set sequentially for rainfall prediction (probabilistic precipitation forecast, QPF). Many scientists have focused their efforts on improving weather forecasting and monitoring, as well as extending QPF technology's ability to predict heavy rainfall events. In recent years, the use of artificial intelligence algorithms for rainfall forecasting has gotten a lot of attention Sankaranarayanan et al.[2019].

The method of rainfall forecasting is a nonlinear system in terms of mathematics. Fuzzy theory and neural network-based artificial intelligence have had a lot of success in dealing with nonlinear technologies in recent years. Adapting these AI algorithms to large-scale rainfall forecasting, on the other hand, is a major challenge. Artificial surveillance programmes include computer systems with high-speed computing and massive data processing capacities, which are difficult to come by. Predictions of rainfall are useful in a variety of fields, including agriculture, construction, tourism, transportation, and health care.

Providing reliable meteorological forecasts to emergency management agencies will help decision-making in the event of a natural disaster. Artificial neural networks (ANNs), which range from simple methods to more advanced techniques like artificial intelligence, are one of the most effective and attractive tools for forecasting activities (AI). In contrast to traditional meteorological strategies, ANNs depend on self-adaptive frameworks that learn from observations and capture operational relationships between data, even when those relationships are ambiguous or difficult to decipher. However, there is a major problem: how to put these ideas into action Wu et al.[2019].

Artificial intelligence operations necessitate computer systems with high-speed computing and massive data processing capacities, which are difficult to come by. In recent years, deep learning has been regarded as a leading mechanism in ANN for completing tasks. Deep learning is a series of multilayer architectures that use unsupervised techniques to learn. The use of unsupervised methods to obtain a lightweight, precise, and non-linear data representation in the hopes of assisting with the prediction task at hand is the main advancement Zhou et al.[2.

This technique has supported computer vision, image recognition, natural language processing, and cognitive science. Deep learning systems including the Restricted Boltzmann Machine (RBM), Conditional RBM, Autoencoder, Recurrent neural network, Convolution and pooling, and Hidden Markov Model have already shown promise in time-series data prediction. Rainfall prediction is a computer programme that predicts how much rain will fall in a given region. It can be done in two ways.The first is to look into the physical laws that affect rainfall, and the second is to devise a strategy for uncovering secret patterns or features that affect physical factors and the processes that lead to them. The second choice is preferred because it requires no mathematical calculations and can be used for complex and non-linear data Zhou et al.[2019];Pham et al.[2020].

Rainfall forecasting has been one of the most critical aspects of most water management systems around the world. One of the most challenging challenges to solve is the lack of rainfall data. The majority of today's rainfall forecasting systems are unable to detect hidden layers or other non-visible features. Forecasts are made using a deep architecture that includes an autoencoder and a multilayer perceptron. The deep architecture's parameters were tweaked, and the resulting structure was examined to see if the conceptual approach was right. According to the findings, the proposed design outperforms the existing state of the art in terms of regular precipitation forecasting accuracy. Section 2 studies background work. Section 3 describes the proposed system model. Section 4 depicts the experimental results followed by conclusion in Section 4.

2. RELATED STUDY

Rainfall is a type of precipitation in which water is scattered over the oceans, environment, and surface of the Earth.Moisture passes through the atmosphere as particulate matter from the Earth's surface, taking heat from the surface with it. This is critical for survival. Rain stores the majority of the world's seawater for plants and animals. When water vapour compresses into cloud droplets, runoff is released into the atmosphere. When considering the various factors involved, the need for rainfall prediction is crucial. National Centers for Environmental Prediction (NCEP), National Weather Service (NWS), and National Oceanic and Atmospheric Administration (NOAA) are examples of climate prediction centres (NOAA), include data on rainfall prediction by evaluating and predicting the impacts of climate change and implying the risks caused for precautionary purposes as well as to minimise economic risks for maximizing profits.

Several research studies have been carried out on rainfall forecasting. Rainfall forecasting data has been available for several years, and the use of neural networks for such time-series data suggests that the computational difficulty is higher, leading to lower precision in predicting the amount of rain in a given region. Deep learning is distinct from neural networks in that the models are altered to increase the number of hidden layers and improve the performance. Since the features in the data are gathered and categorised as a single classifier, deep-learning models have the benefit of implementing a single algorithm.

For forecasting cumulative daily precipitation for the next day, a Deep Learning-based framework is suggested. To mitigate and obtain non-linear correlations between categories, a multilayer perceptron is used for forecasting, and an autoencoder is used to mitigate and capture non-linear correlations between attributes. In comparison to previous plans, this design shows an increase in the ability to forecast cumulative daily precipitation for the next day Hernandez et al.[2016]. For countries like India, whose economy is heavily dependent on agriculture, the accuracy of rainfall forecasting is crucial. Statistical methods struggle to provide sufficient precision for rainfall forecasting due to the complex nature of the atmosphere. Because of the nonlinearity of rainfall data, Artificial Neural Networks are a superior technique. In a tabular format, the review work and comparison of various methods and algorithms used by researchers for rainfall prediction are shown. The aim of this paper is to provide non-experts with easy access to rainfall prediction techniques and approaches Parmar et al.[2017].

The use of Deep Learning Architectures (LSTM and ConvNet) to obtain models of rainfall precipitation is proposed, with LSTM having an RMSE of 2.55 and ConvNet having an RMSE of 2.44, stating that Deep Learning models would be accurate and reliable for modellers for any time series dataset Aswin et al.[2018]. By analysing temporal statistical models such as long short-term memory (LSTM) and temporal convolutional networks (TCN), a novel lightweight data-driven weather forecasting system is developed, and its results are compared to conventional analytical machine learning methods, statistical modelling strategies, and a dynamic ensemble method, as well as some well-weather analysis and prediction. Multi-input multi-output and multi-input single-output regressions are used to assess the system deep learning networks with LSTM and TCN layers. The study demonstrates that the lightweight model suggested outperforms the commonly used model Hewage et al.[2021].

The Rainfall Nowcasting Network (RN-Net) is a dual-input dual-encoder recurrent neural network. It forecasts grid rainfall data for the next two hours using past grid rainfall data curve fitting by automated weather stations and doppler radar mosaic data as raw data. On the Southeastern China dataset, we keep track of how far we've come. The RN-rainfall Net's nowcasting hazard scores have exceeded 0.523, 0.503, and 0.435 in 0.5 h, 1 h, and 2 h, respectively, using a 0.25 mm threshold. The hazard scores have risen by about four times, three times, and three times, respectively, as opposed to the Weather Analysis and Forecasting model rainfall nowcasting Zhang et al.[2021].

For the forecasting of summer monsoon rainfall, numerous deep learning methods such as one-dimensional Convolutional Neural Networks, Multi-layer Perceptrons, and Large Deep Neural Networks are being researched and tested in the Indian state of Rajasthan. Two different types of datasets are analysed for the experiments. Rainfall data for 484 coordinates within the geographical borders of Rajasthan is selected from the IMD gridded dataset. The water management department obtained rainfall data from 158 rain gauge stations. On each of these data sets, this paper compares various algorithms, and it is revealed that the Deep Wide Neural Network-based significantly achieves the others. On each of these data sets, this paper compares various algorithms and concludes that the Deep Wide Neural Network-based model outperforms the other different objectives Bajpai et al.[2021].

Deep learning has recently proven to be an effective method for solving complex problems and analysing large amounts of data. The proposed framework used generative adversarial networks to create a rainfall prediction system that analysed India's rainfall data and predicted future rainfall. In the conceptual methodology, a GAN network was used as a predictor, with an LSTM network algorithm as a discriminator and a convolution neural network model as a classifier. The LSTM algorithm is well adapted to predicting time series results like rainfall. The results of the experiments show that the proposed approach correctly predicts the outcomes with a rate of 99 percent accuracy. Rainfall forecasting aids farmers in cultivating their crops.Raj et al.[2021].

Forecasting atmospheric phenomena such as rainfall can help to avoid disasters, given the effect of weather on social and economic activities.Several approaches to dealing with this aim have been suggested in recent years. There has been a lot of research done in the field of rainfall prediction, and one can see how effective neural networks are at solving the problem Schultz et al.[2021];Sanekane et al.[2021]. Artificial neural networks (ANNs) are one of the expert machines that has been used more for weather prediction studies, according to the literature review Naik et al.[2020].In particular, when dealing with different applications and anomalies, such as meteorological phenomena, neural networks are an important mechanism. Deep Learning is also being hailed as a promising new tool for forecasting weather variables.

3.PROPOSED SYSTEM MODEL

Time-series modelling has long been used in academic research in fields such as climate modelling, biomedical sciences and healthcare, and corporate decision in retail and financing, to name a few. Classic approaches to learning temporal aspects, such as autoregressive (AR), simple exponential, or structural time-series models, have concentrated on parametric models guided by domain expertise, while modern machine learning techniques allow for data-driven learning of temporal dynamics. Deep learning has become an important part of the next generation of time-series forecasting as data access and computing capacity have improved in recent years.Figure 1 depicts the framework of proposed rainfall prediction.

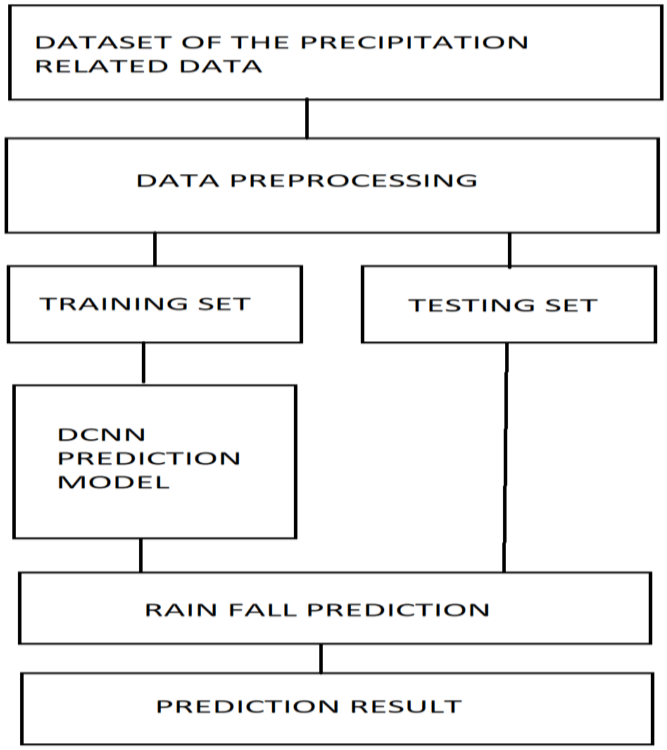

Figure 1. Proposed framework of rainfall prediction

3.1 DATASET

The problem discussed in this study is estimating total rainfall precipitation for the following day based on data obtained the day before. We specifically used a data set from kaggle.com.This dataset includes regular weather observations from a number of Australian weather stations dating back about ten years. The vector to forecast is RainTomorrow.

Table 1. Dataset Description

Feature	**Description**
Date	Date of the observation
Location	The location of the weather station's common name
MinTemp	The lowest temperature, measured in degrees Celsius.
MaxTemp	In degrees Celsius, the mean temperature
Rainfall	In mm, the amount of rain that fell that day.
Evaporation	In the 24 hours leading up to 9 a.m., the so-called Class A pan evaporation (mm) was measured.
Sunshine	The number of hours in the day when the sun is shining brightly.
WindGustDir	In the 24 hours leading up to midnight, the path of the highest wind gust
WindGustSpeed	The fastest wind gust's intensity (km/h) in the 24 hours leading up to midnight
WindDir9am	At 9 a.m., the wind was blowing in the opposite direction.

Preprocessed data has been stored in the data warehouse. In order to achieve data integrity and standardisation, it has undergone an ETL (Extract, Transform, and Load) phase.

3.2 DATA PRE-PROCESSING

3.2.1 Data cleansing

The data centre has been used to store preprocessed data. In particular, it went through an ETL (Extract, Transform, and Load) process. Data cleaning is the method of locating and correcting inaccurate or outlier records in a dataset, as well as replacing, updating, or deleting inaccurate data that could impact the model's accuracy. There were a few days in our case when the required factors were not recorded, and if there was trace precipitation, the rainfall in centimetres was marked as ‘T'. We can't deal with alphabets in the data because the algorithm needs numbers. As a result, we must clean the data before using it in our model.

3.2.2 Finding Correlation and Correlation coefficient

Correlation is a statistical indicator that indicates the degree to which two or more variables are interdependent. A positive correlation indicates that when one factor rises, the other rises with it, while a negative correlation means that when one factor rises, the other falls. A correlation coefficient is a statistical measure that can be used to decide which attributes in a dataset are more linked to or dependent on the output variable, allowing for easier feature selection.

3.2.3 Normalization (Scaling of Data)

To normalise the range of individual variables or data features, the scaling or normalisation approach is used.Data normalisation is normally done as part of the preprocessing stage. Since all of the values are translated between a certain set of values, normalising the data makes the model less complex. In our case, we used a range of -1 to 1 to normalise the data.

3.3 DEEP CONVOLUTIONAL NEURAL NETWORK

Convolutional Neural Networks (ConvNet) are a type of Artificial Neural Network (ANN) with balanced and distorted neurons. Each neuron receives the inputs and then conducts a non-linear input feature operation. The ConvNet Framework, unlike other neural networks, can handle multi-dimensional data. Convolution with Non-Linearity (ReLU), Pooling, and Fully Connected Layer are the three strategies used in this layout in the range of separate surfaces.

3.3.1 Convolutional Layer

The main aim of a Convolutional Layer is to retrieve data features. Convolution helps in the correlation of image pixels by expressing the features of a matrix narrower than the data. The Feature Map or Convolved Feature is the output of the filtered input data, and the Feature Map or Convolved Feature is the output of the Feature Map or Convolved Feature. Three variables control the size of this layer: width, stride, and zero-padding. The Rectified Linear Unit (ReLU) is a non-linear operation that substitutes all negative pixel values in a convolutional layer's output with zero on an element-by-element basis. Since most real-time data necessitates non-linear learning, this function's goal is to incorporate non-linearity into ConvNet.This layer's performance is referred to as the 'Rectified' function diagram.

3.3.2 Max-Pooling Layer

Pooling holds the most critical info while reducing the size of each convolutional layer's output. Max Pooling is a spatial neighbourhood that uses the rectified feature map's highest values. Because of the reduction in size and processing, the issue of computational burden is managed.

3.3.3 Fully-Connected (FC) Layer

The output layer for this layer is an MLP with a Softmax function, which means that all of the neurons on this layer are related. The aim of this layer is to forecast input test data from train data using features. The probability of a cumulative production of 1 is always 1.

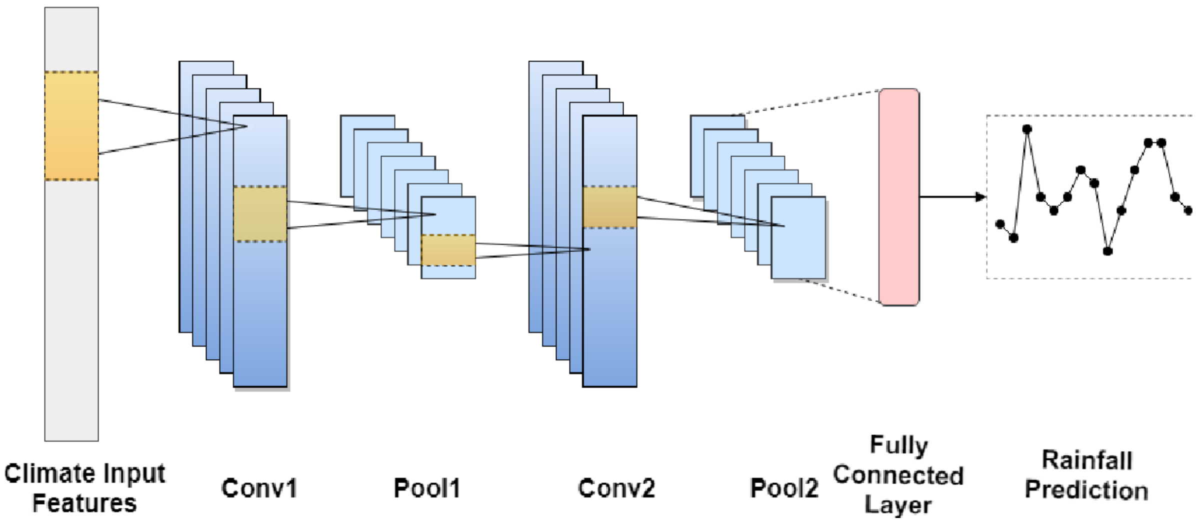

Figure 2. Proposed ConvNet Framework

The Neural System is designed with specific attributes such as humidity, temperature, and pressure, and it continues to recognise and perceive rainfall feature vectors using the performance of the training dataset. One input neuron, one secret neuron, and one output neuron could make up a very simple neural network. It increases the weights of the components of multiple dependent variables = input parameters before passing them through a ReLU activation function and a unit phase function.

oj = f (Σ wi, jai + bi) (1)

Input Layer

The amount of neurons in the input layer will be the same as the number of input functions. The inputs will be accumulated by weights before being sent to the secret layer for processing.

An activation function(Rectified Linear Unit ReLu)

Rectified linear Unit is the abbreviation for Rectified linear Unit. It's commonly used in deep learning. Only a portion of the ReLU activation mechanism has been corrected. When z is less than zero, f(z) equals 0. All negative values are converted to zero when z is positive or zero, and f(z) equals z.

The activation function can be denoted as:

R(z)=Max (0 to infinity) (2)

Range: [0 to infinity] (3)

When a neural network is trained on a training set, it is given a set of weights. During the training cycle, weights are optimised and optimal weights are generated. The neuron generates a weighted sum of inputs, as shown below.

Y = Σ (weight * input) + bias (4)

Consider the following scenario: x1, x2,... xn are the inputs, and w1, w2,... wn are the weights.

Finally, to prepare the output, the computed value is transferred to the activation function, which in our case is ReLU.

Activation function (x1w1 + x2w2 + ….. + xnwn + bias) (5)

Output layer

To forecast only one variable, Rainfall, only one output layer is used.

3.4 TESTING AND TRAINING

With percentages of 70, 15, and 15%, the dataset is divided into preparation, validation, and testing. As a result, 70% of the total samples that matched the dataset were chosen at random as preparation. Just 15% of the remaining samples were chosen for validation, while the remaining 15% were held for processing.

4. EVALUATION AND EXPERIMENTAL RESULTS

We used Python 3 and Google Colab's Jupyter Notebook for all of our experiments and classifier development. Sckit Read, Matplotlib, Seaborn, Pandas, Numpy, and Imblearn were among the libraries we used. Weka was used to implement the Decision Table. We ran experiments with different input data, beginning with the initial dataset, then the undersampled dataset, and finally the oversampled dataset.

Performance Metrics

Chance of possibilities

A probability of precipitation (POP), also known as chance of precipitation or chance of rain, is a measure of the likelihood of at least some minimal amount of precipitation occurring during a given forecast period and region.

Amount of rainfall

The rainfall amounts shown here help to explain how rainfall will vary depending on the type of weather in a given time span. Both numbers refer to the likelihood of getting at least that much rain.

We computed the output after feeding multiple inputs into our deep convolutional Neural Networks, such as temperature, humidity, and wind speed. Figure 3 depicts the comparative study between the current ANN and the proposed DeepConvNet.

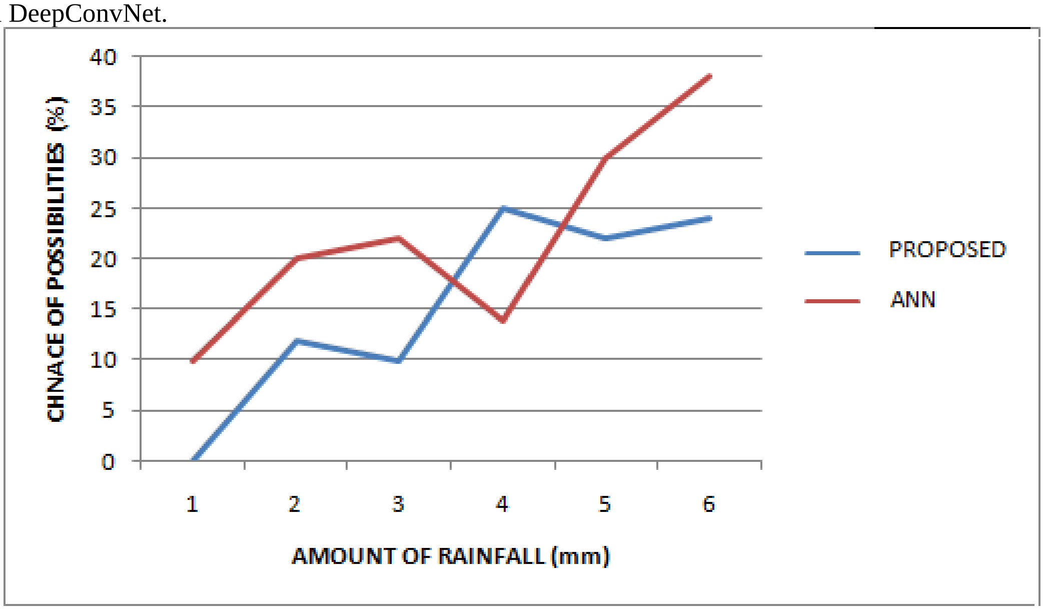

Figure 3. Comparison analysis Proposed DeepConvNet Vs CNN

The vertical axis represents the likelihood of our algorithm's forecasts, while the horizontal axis represents the amount of rainfall in real time. A perfect prediction is represented by a straight line with a slope of one and no intercept. According to the above analysis, the proposed approach outperforms the conventional method.

5. CONCLUSION

Rainfall is one of the most important aspects of agriculture, and forecasting it is one of the most difficult tasks. Weather and rainfall are nonlinear and complex phenomena that require sophisticated computer modelling and simulation to accurately forecast. By allowing self-learning data labels, Deep Learning has recently allowed the development of a data-driven model for a time series dataset. It allows for the detection of anomalies/changes in time series data as well as the prediction of future event data using historical data.The work's primary goal is to obtain a better model of the rainfall precipitation by using Deep Learning Architectures and determining the better architecture with convolutional neural network for any time series dataset, Deep Learning models will be effective and efficient for the modellers.

REFERENCES

1. Aguasca, Ricardo & Nieves, Dagoberto & Méndez, Máximo. (2019). Comparative Analysis of Rainfall Prediction Models Using Machine Learning in Islands with Complex Orography: Tenerife Island. Applied Sciences. 9. 4931. 10.3390/app9224931.
2. Aswin, S & Srikanth, Dr.geetha & Ravi, Vinayakumar. (2018). Deep Learning Models for the Prediction of Rainfall. 0657-0661. 10.1109/ICCSP.2018.8523829.

3. Bajpai, Vikas & Bansal, Anukriti. (2021). A Deep and Wide Neural Network-based Model for Rajasthan Summer Monsoon Rainfall (RSMR) Prediction.
4. Hernandez, Emilcy & Sanchez-Anguix, Víctor & Julián, Vicente & Palanca, J. & Duque, Néstor. (2016). Rainfall Prediction: A Deep Learning Approach. 151-162. 10.1007/978-3-319-32034-2_13.
5. Hewage, Pradeep & Trovati, Marcello & Pereira, Ella & Behera, Ardhendu. (2021). Deep learning-based effective fine-grained weather forecasting model. Pattern Analysis and Applications. 24. 10.1007/s10044-020-00898-1.
6. Naik, Akshay & Deorankar, A. & Ambhore, Premchand. (2020). Rainfall Prediction based on Deep Neural Network: A Review. 98-101. 10.1109/ICIMIA48430.2020.9074892.
7. Parmar, Aakash & Mistree, Kinjal & Sompura, Mithila. (2017). Machine Learning Techniques For Rainfall Prediction: A Review. 2017 International Conference on Innovations in information Embedded and Communication Systems (ICIIECS).
8. Pham, Binh & Le, Lu & Le, Tien-Thinh & Bui, Kien-Trinh & Le, Vuong & Ly, Hai-Bang & Prakash, Indra. (2020). Development of advanced artificial intelligence models for daily rainfall prediction. Atmospheric Research. 237. 10.1016/j.atmosres.2020.104845.
9. Raj, Venkat & Balasubramanian, C. & Kaliappan.,M.E.,Ph.D, Dr.M. (2021). Rainfall prediction using generative adversarial networks with convolutional neural networks. Soft Computing. 25. 1-14. 10.1007/s00500-020-05480-9.
10. Sankaranarayanan, Suresh & Prabhakar, Malavika & Satish, Sreesta & Jain, Prerna & Ramprasad, Anjali & Krishnan, Aiswarya. (2019). Flood prediction based on weather parameters using deep learning. Journal of Water and Climate Change. 11. 10.2166/wcc.2019.321.
11. Senekane, Makhamisa & Mafu, Mhlambululi & Taele, Benedict. (2021). Weather Nowcasting Using Deep Learning Techniques. 10.5772/intechopen.84552.
12. Schultz, Martin & Betancourt, Clara & Gong, Bing & Kleinert, Felix & Langguth, Michael & Leufen, Lukas & Mozaffari, Amirpasha & Stadtler, Scarlet. (2021). Can deep learning beat numerical weather prediction?. Philosophical Transactions of The Royal Society A Mathematical Physical and Engineering Sciences. 379. 10.1098/rsta.2020.0097.
13. Wu, Zhengdong & Wu, Fujian & Chai, Jin & Zhan, Choujun & Yu, Zhiwen. (2019). Prediction of Daily Precipitation Based on Deep Learning and Broad Learning Techniques. 513-519. 10.1109/ISKE47853.2019.9170361.
14. Zhang, Fuhan & Wang, Xiaodong & Guan, Jiping & Wu, Meihan & Guo, Lina. (2021). RN-Net: A Deep Learning Approach to 0–2 h Rainfall Nowcasting Based on Radar and Automatic Weather Station Data. Sensors. 21. 1981. 10.3390/s21061981.
15. Zhou, Kanghui & Zheng, Yongguang & Li, Bo & Dong, Wansheng & Zhang, Xiaoling. (2019). Forecasting Different Types of Convective Weather: A Deep Learning Approach. Journal of Meteorological Research. 33. 797-809. 10.1007/s13351-019-8162-6.